Sharp Indistinguishability Bounds from Non-Uniform Approximations

Christopher Williamson
The SW7 Group, Hong Kong, China

Abstract

We study the basic problem of distinguishing between two symmetric probability distributions over n bits by observing k bits of a sample, subject to the constraint that all $(k-1)$-wise marginal distributions of the two distributions are identical to each other. Previous works of Bogdanov et al. [3] and of Huang and Viola [8] have established approximately tight results on the maximal possible statistical distance between the k-wise marginals of such distributions when k is at most a small constant fraction of n. Naor and Shamir [12] gave a tight bound for all k in the special case $k = n$ and when distinguishing with the OR function; they also derived a non-tight result for general k and n. Krause and Simon [9] gave improved upper and lower bounds for general k and n when distinguishing with the OR function, but these bounds are exponentially far apart when $k = \Omega(n)$. In this work we provide sharp upper and lower bounds on the maximal statistical distance that hold for all k and n. Upper bounds on the statistical distance have typically been obtained by providing *uniform* low-degree polynomial approximations to certain higher-degree polynomials. This is the first work to construct suitable non-uniform approximations for this purpose; the sharpness and wider applicability of our result stems from this non-uniformity.

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization

Keywords and phrases bounded indistinguishability, randomness, secret sharing, polynomial approximation

Digital Object Identifier 10.4230/LIPIcs.STACS.2022.59

1 Introduction

We consider pairs of distributions μ and ν over $\{0,1\}^n$. The distributions μ and ν are said to be perfectly j-wise indistinguishable if for any subset $S \subseteq [n]$ of size at most j, the marginal distributions μ_S and ν_S over indices in S are identically distributed. The distributions are k-wise reconstructible with advantage ϵ (alternatively, ϵ-distinguishable) if there exists a set $S \subseteq [n]$ of indices of size k and a statistical test $T : \{0,1\}^{|S|} \to \{0,1\}$ such that

$$|\mathbb{E}_{X\sim\mu}[T(X|_S)] - \mathbb{E}_{Y\sim\nu}[T(Y|_S)]| \geq \epsilon,$$

where $X|_S$ is the restriction of random variable X to the bits located at the indices in S. Equivalently, μ and ν are j-wise indistinguishable if all size $\leq j$ marginal distributions have 0 total variation distance; μ and ν are k-wise reconstructible with advantage ϵ if any of the k-wise marginals have total variation distance at least ϵ. The distributions are symmetric if μ and ν are invariant under permutation (see definitions in Section 2); for such distributions the size of S is relevant for distinguishing but not the choice of indices.

Cryptographic motivation

Work of Bogdanov et al. [2] considered this notion of indistinguishability as a way to capture cryptographic secret sharing schemes in a minimal setting. Their observation was that a single bit secret can be shared by sampling n bits from μ or from ν, depending on the secret: the j-wise indistinguishability of the distributions provides a security guarantee that any size $\leq j$ coalition of colluding parties learn nothing about the secret from their joint

shares. The secret reconstruction function for the scheme is a test T applied over the shares of sufficiently many (possibly all) parties. A key question in their work was how large j could be taken so that there exists a T that is both computable with AC^0 circuits and has reconstruction advantage $\epsilon = \Omega(1)$ against some pair of j-wise indistinguishable distributions. The special case of $T = \mathsf{OR}$ captures the notion of visual cryptography, introduced by Naor and Shamir [12] in 1994. Their work considered j-wise indistinguishable distributions and attempted to maximise the reconstruction advantage of OR taken over $j + 1$ bits. They gave a tight bound for all j in the special case $j = n - 1$; they also derived a non-tight result for general j and n. Krause and Simon [9] gave improved upper and lower bounds for general j and n, but these bounds are exponentially far apart when $j = \Omega(n)$. In this work (as in [4], [8]), we consider the statistical distance between the distributions, which includes the study of tests T that are not in AC^0 and reconstruction advantage ϵ that may be vanishing. These results can be interpreted as tight existence or non-existence conditions for 1-bit secret sharing schemes of arbitrary reconstruction complexity.

Approximate degree motivation

The work [2] largely proceeded by a connection to the theory of approximate degree of Boolean functions. The ϵ-approximate degree of a Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$, denoted $\widetilde{\deg}_\epsilon(f)$, is the least degree of a multivariate real-valued polynomial p such that $|p(x) - f(x)| \leq \epsilon$ for all inputs $x \in \{0,1\}^n$. This quantity has received significant attention, owing to its polynomial equivalence to many other complexity measures including sensitivity, exact degree, deterministic and randomized query complexity [14], and quantum query complexity [5]. By linear programming duality, f has ϵ-approximate degree more than j if and only if there exists a pair of probability distributions μ and ν over $\{0,1\}^n$ such that μ and ν are j-wise indistinguishable and n-wise reconstructible with advantage 2ϵ by f (see for example [2, 17, 6]). The approximate degree of all symmetric Boolean functions was resolved in the constant-error regime $\epsilon = \Theta(1)$ by Paturi [15] and in the general error regime by de Wolf [7] using an argument based on quantum algorithms (see also Sherstov [18] and Bun and Thaler [6]). This implies tight upper and lower bounds on the ability of any given symmetric Boolean function to reconstruct from indistinguishable distributions when given access to a full sample of n bits. In this work, we consider reconstruction when given access to a subset of the bits.

Prior work

Works of Bogdanov et al. [3] and of Huang and Viola [8] extended the study of the indistinguishability of symmetric distributions to the setting of reconstructing with a subset of indices. They considered the extent to which symmetric j-wise indistinguishable distributions must have statistically close k-wise marginals for $k > j$. In particular, [3] shows that if μ and ν are symmetric over n-bit strings and perfectly j-wise indistinguishable, then the statistical distance between k-wise marginals is at most $O(j^{3/2}) \cdot e^{-j^2/1156k}$ for all $j < k \leq n/64$. The analogous result in [8] gives a similar bound and also applies to k at most some (unspecified) constant fraction of n. A matching lower bound given in [3] shows that there exists a pair of distributions that are j-wise indistinguishable but reconstructable with the OR_k function with advantage at least $k^{-1/2} \cdot e^{-O(-j^2/k)}$. This lower bound extends to all $j < k \leq n$.

We note that in the sharp reconstruction setting where $k = j + 1$, the upper bound $O(j^{3/2}) \cdot e^{-j^2/1156k}$ solely demonstrates a behaviour of monotonic exponential decrease in k. Extension of this upper bound to $k > n/64$ is of particular interest because it would illuminate

and quantify the behaviour that the maximum statistical distance must reach a minimum at some $k \in [n/64, n]$ and start to increase once k becomes large enough. Specifically, we know this behaviour changes because the XOR_n function can distinguish perfectly between a pair of symmetric distributions that are $(n-1)$-wise indistinguishable (the two distributions are uniform over n bits, conditioned on the sum of all n bits being either even or odd).

In addition, the lower bound for OR in [3] does apply to k up to n but is unlikely to come close to matching any upper bound on statistical distance since OR is a weak statistical distinguisher and the bound is also monotonically decreasing.

Our contribution

In the present work, we extend the results of [3] and [8] to the setting of parameters where k may range freely from 0 to $n-1$ (the case $k = n$ is trivial in light of the XOR example above). We consider the sharp threshold reconstruction setting $j = k - 1$ (see prior work section) and our results are tight up to polynomial factors.

▶ **Theorem 1.** *There exists an absolute constant c such that for any $k \in [0, n-1]$ and any pair of symmetric $(k-1)$-wise indistinguishable distributions μ, ν over $\{0,1\}^n$, the statistical distance between all k-wise marginals of μ and ν is at most:*

$$O(n^c) \cdot \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}$$

▶ **Theorem 2.** *For any $k \in [0, n-1]$, there exists a statistical test $T : \{0,1\}^k \to \{0,1\}$ and a pair of symmetric $(k-1)$-wise indistinguishable distributions μ, ν such that the reconstruction advantage of T is at least*

$$\frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}.$$

An indistinguishability game. The upper and lower bounds of Theorems 1 and 2 allow us to approximately find the value of k so that the task of using k bits to distinguish symmetric perfectly $(k-1)$-wise indistinguishable distributions over $\{0,1\}^n$ is most difficult. For n an integer fixed in advance, this can be expressed in terms of a game:

- Player 1 selects integer $k < n$.
- Player 2 chooses a pair μ, ν of symmetric, perfectly $(k-1)$-wise indistinguishable distributions over $\{0,1\}^n$.
- The payoff p_2 of Player 2 is defined as the statistical distance between the k-wise marginals of μ and ν and the payoff of Player 1 is $p_1 = 1 - p_2$.

We show that the optimal strategy of Player 1 is essentially to select $k = 0.6n$. More precisely:

▶ **Corollary 3.** *In the indistinguishability game defined above, let k^* be an optimal strategy of Player 1. Then, for any arbitrarily small positive constant ϵ, and n sufficiently large, k^* satisfies:*

$$|k^* - 0.6n| \leq \epsilon n.$$

A naive assumption would be that for a fixed n and k ranging in $[0, n]$, the behaviour of the quantity being bound in Theorems 1 and 2 is symmetric around the centre of the interval $[0, n]$, or namely that the optimal strategy of Player 1 is to choose $k = n/2$. Corollary 3 demonstrates that this is false.

Finally, Theorem 1 can be translated into Fourier analytic language. We consider a function $f : \{-1,1\}^n \to \mathbb{R}$ and write it in the Fourier basis as $f(x) = \sum_{S\subseteq[n]} \hat{f}(S) \cdot \prod_{i\in S} x_i$, where $\hat{f}(S) = \mathbb{E}_{x\in\{-1,1\}^n}[f(x) \cdot \prod_{i\in S} x_i]$.

▶ **Corollary 4.** *Let $f : \{-1,1\}^n \to \mathbb{R}$ be a real-value Boolean function that is symmetric ($|S| = |S'| \implies \hat{f}(S) = \hat{f}(S')$), has no low-degree terms ($|S| \leq k-1 \implies \hat{f}(S) = 0$), and has low 1-norm ($\sum_{z\in\{-1,1\}^n} |f(z)| = 1$). Then, for any S of size k, we have:*

$$\hat{f}(S) \leq O(n^c) \cdot \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}.$$

Techniques and roadmap

The established technique to develop indistinguishability upper bounds for symmetric distributions is to decompose an arbitrary statistical test into a small basis using the fact that without loss of generality, the best test is a symmetric function. The basis we work over is Q_w for $w = 0, ..., k$ where Q_w is a Boolean function that observes k bits and accepts if and only if the observed Hamming weight is exactly w. Providing a low-degree polynomial approximation to Q_w rules out the existence of distributions that can be reconstructed with Q_w. In practice, a Minsky-Papert symmetrization (see Fact 13 for details) is applied to Boolean function Q_w to reduce the problem of its approximation to a problem of approximating a real-valued univariate polynomial with a lower degree polynomial. Due to the discrete domain of Q_w (the Boolean cube), the univariate approximations need not be uniform, but are instead over a set of separated points on the real line, each representing a Hamming weight of input (for example, $-1, -1+2/n, ..., 1-2/n, 1$). However, this fact is not typically exploited, and previous works have constructed approximations that have low error over the entire interval $[-1, 1]$.

Prior works have not obtained statistical distance upper bounds for k close to n because the approach taken to approximating (the symmetrized version of) Q_w has been to use Chebyshev polynomials to provide uniform approximations to Q_w over all of $[-1, 1]$ instead of discrete approximations. This strategy breaks down for large k because the difficulty of uniform approximations diverges from the difficulty of the approximation over the relevant discrete set. This observation motivates the use of discrete Chebyshev polynomials (also known as Gram polynomials) to construct approximations that yield upper bounds on the maximum statistical distance.

We provide a lower bound on statistical distance based on hardness of approximation with discrete Chebyshev polynomials. This follows from their orthogonality and from linear programming duality. We believe that prior techniques via orthogonality of (non-discrete) Chebyshev polynomials could be used to show this result (indeed the lower bound result from [3] applies to all k up to n and the technique they use should be extendable to distinguishers more powerful than OR).

2 Preliminaries

We use the standard notion of statistical distance (total variation distance) throughout this paper. Specifically, the statistical distance between μ and ν is $\frac{1}{2}||\mu - \nu||_1$, or $\frac{1}{2}\sum_{z\in\{0,1\}^n} |\Pr_{X\sim\mu}[X = z] - \Pr_{Y\sim\nu}[Y = z]|$. This is equivalent to the reconstruction advantage of the optimal statistical test for μ and ν that observes all n bits.

We will be working with polynomial approximations over different discrete sets of points. We define D_n^{out} as the set of points $\{-1, -1+2/n, ..., 1\}$ and D_n^{in} as $\{-1+1/n, -1+3/n, ..., 1-1/n\}$. It is easy to check that $|D_n^{\text{out}}| = n+1$, that $|D_n^{\text{in}}| = n$. Further, we have the basic relationships:

$$D_n^{\text{out}} = \left\{ \frac{n+1}{n} \cdot x : x \in D_{n+1}^{\text{in}} \right\} \tag{1}$$

$$D_n^{\text{in}} \subset \left\{ x - \frac{1}{n} : x \in D_n^{\text{out}} \right\} \tag{2}$$

For simplicity we will use $\lesssim$ to hide factors polynomial in n.

Symmetric distributions and functions

Let $f : \{0,1\}^n \to \mathbb{R}$ be a function. We say that f is symmetric if the output of f depends only on the Hamming weight of its input. A probability distribution μ is symmetric if the corresponding probability mass function mapping inputs to probabilities is a symmetric function. We also will need two further facts about distinguishing symmetric distributions. Proofs of these appear in [3].

▶ **Fact 5.** *Suppose that μ is a symmetric distribution over $\{0,1\}^n$. For $S \subseteq \{0, ..., n\}$, let $\mu|_S$ denote the restriction of μ to the indices in S. Then, $\mu|_S$ is also symmetric.*

▶ **Fact 6.** *Suppose that μ and ν are symmetric distributions over $\{0,1\}^n$. Then without loss of generality, the best statistical test $Q : \{0,1\}^n \to [0,1]$ for distinguishing between μ and ν is a symmetric function. In particular, we have:*

$$\max_{\text{symmetric } Q} \{\mathbb{E}_{X\sim\mu}[Q(X)] - \mathbb{E}_{Y\sim\nu}[Q(Y)]\} = \max_{Q} \{\mathbb{E}_{X\sim\mu}[Q(X)] - \mathbb{E}_{Y\sim\nu}[Q(Y)]\}.$$

2.1 Discrete Chebyshev polynomials

The discrete Chebyshev polynomials, for parameter n are a family of real polynomials $\{\phi_d\}_{d=0,...,n-1}$. Borrowing notation from [1], we have that the polynomials have the following properties:

- The family of polynomials $\{\phi_d\}_{d=0,...,n-1}$ are orthogonal with respect to the bilinear form given by

$$(\phi_i, \phi_j) := \frac{1}{n} \cdot \sum_{x \in D_n^{\text{in}}} \phi_i(x) \cdot \phi_j(x) \tag{3}$$

- For each d:

$$||\phi_d|| := (\phi_d, \phi_d)^{1/2} = 1 \tag{4}$$

- For each d:

$$\deg(\phi_d) = d \tag{5}$$

- The polynomials satisfy the recurrence:

$$\phi_d(x) = 2\alpha_{d-1} \cdot x \cdot \phi_{d-1}(x) - \frac{\alpha_{d-1}}{\alpha_{d-2}} \cdot \phi_{d-2}(x) \tag{6}$$

$$\alpha_{d-1} = \frac{n}{d} \cdot \left(\frac{d^2 - 1/4}{n^2 - d^2}\right)^{1/2} \tag{7}$$

where we have $\phi_0 = 1$, $\phi_{-1} = 0$, and $\alpha_{-1} = 1$.

Every degree $k < n$ polynomial $p\colon \mathbb{R} \to \mathbb{R}$ has a unique expansion in the discrete Chebyshev basis:

$$p(t) = \sum_{d=0}^{k} c_d \phi_d(t),$$

where $c_0, \ldots, c_k$ are the *discrete Chebyshev coefficients* of p.

2.2 Bounds on factorials and binomial coefficients

We will make use of double factorials, which are given by:

$$n!! := \prod_{i=0}^{\lfloor n/2 \rfloor} n - 2i \tag{8}$$

and satisfy, when n is even:

$$n!! = 2^{n/2} \cdot (n/2)! \tag{9}$$

For n odd, we simply observe that $(n-1)!! \lesssim n!! \lesssim (n+1)!!$ and apply the bound in (9).

We will bound factorials using

$$n! = \Theta(\sqrt{n}) \cdot \left(\frac{n}{e}\right)^n \tag{10}$$

and the central binomial coefficient using

$$\binom{2n}{n} = \Theta(1/\sqrt{n}) \cdot 2^{2n} \tag{11}$$

2.3 Approximate degree of Boolean functions

Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. We will use $\widetilde{\deg}_\epsilon(f)$ to denote the minimum degree of any real polynomial $p : \{0,1\}^n \to \mathbb{R}$ that approximates f to within ϵ at every point in $\{0,1\}^n$.

Paper organisation. In Section 3, we prove a lemma about the expression of monomials in the discrete Chebyshev basis. In Section 4 we construct discrete approximations to the monomial and prove a complementary hardness of approximation result in Section 5. In Section 6 we quantify the precise approximation problem that we need to solve and justify this using symmetrization and linear programming duality techniques. Sections 7 and 8 justify Theorem 1 and Theorem 2, respectively. In the Appendix, we handle proofs of some technical claims and proofs of Corollaries 3 and 4.

3 Monomials in the discrete Chebyshev basis

▶ **Lemma 7.** *Fix integer $k < n$. Let C be the leading coefficient in the expansion of x^k in the discrete Chebyshev polynomial basis with parameter n. Then, C satisfies:*

$$\frac{1}{(2n)^k} \cdot \sqrt{\frac{(n+k)!}{(n-k)!}} \lesssim C \lesssim \frac{1}{(2n)^k} \cdot \sqrt{\frac{(n+k)!}{(n-k)!}}$$

Proof. From the recurrence definition of the discrete Chebyshev polynomials in Equation 6, we have that

$$x \cdot \phi_i = \frac{1}{2\alpha_{i-1}}\phi_{i+1} + \frac{1}{2\alpha_{i-2}}\phi_{i-1}$$

Application of this recursion, along with the base case $\phi_0 = 1$, yields that the discrete Chebyshev representation of x^k has its highest degree coefficient given by

$$C = \prod_{i=0}^{k-1} \frac{1}{2\alpha_i} = 2^{-k} \cdot \prod_{i=0}^{k-1} \frac{1}{\alpha_i} \tag{12}$$

From the definition of the α_i in Equation 7, we have that

$$\begin{aligned}
\prod_{i=0}^{k-1} \alpha_i &= \frac{n^{k-1}}{(k-1)!}\sqrt{\frac{(1^2-1/4)(2^2-1/4)\cdot\ldots\cdot((k-1)^2-1/4)}{(n^2-1^2)(n^2-2^2)\cdot\ldots\cdot(n^2-(k-1)^2)}} \\
&= \frac{n^{k-1}}{(k-1)!}\sqrt{\frac{(1^2-1/4)(2^2-1/4)\cdot\ldots\cdot((k-1)^2-1/4)}{(n+1)(n-1)(n+2)(n-2)\cdot\ldots\cdot(n+k-1)(n-k+1)}} \\
&= n^{k-1} \cdot \sqrt{\frac{n!(n-k)!}{(n-1)!(n+k-1)!}} \cdot \sqrt{\frac{(1^2-1/4)(2^2-1/4)\cdot\ldots\cdot((k-1)^2-1/4)}{(k-1)^2\cdot\ldots\cdot 1^2}} \\
&= n^{k-1} \cdot \sqrt{\frac{n\cdot(n-k)!}{(n+k-1)!}} \cdot \sqrt{\prod_{i=1}^{k-1}\frac{(i-1/2)(i+1/2)}{i^2}}.
\end{aligned}$$

Application of the upper bound in Claim 20 yields that $\prod_{i=0}^{k-1}\alpha_i \leq n^{k-1}\cdot\sqrt{\frac{n\cdot(n-k)!}{(n+k-1)!}}$. This, in conjunction with Equation 12, justifies the lower bound in the statement of this lemma. For the upper bound, application of the lower bound in Claim 20 yields that $\prod_{i=0}^{k-1}\alpha_i \geq \frac{3}{4(k-1)^2}\cdot n^{k-1}\cdot\sqrt{\frac{n\cdot(n-k)!}{(n+k-1)!}}$, which completes the proof in light of Equation 12. ◀

4 Discrete approximations for the monomial

▶ **Corollary 8.** *Fix integers $k < n$. There exists a degree at most $k-1$ polynomial approximation for the monomial x^k over D_n^{out} with error ϵ satisfying:*

$$\epsilon \lesssim (2(n+1))^{-k} \cdot \sqrt{\frac{(n+k+1)!}{(n-k+1)!}}$$

Proof. It suffices to provide a degree at most $k-1$ approximation p to the monomial $(\frac{n+1}{n}\cdot x)^k$ over D_{n+1}^{in} because then the degree at most $k-1$ approximation $p' := p(\frac{n}{n+1}\cdot x)$ will be an approximation to x^k over D_n^{out}, by Equation 1. By Lemma 7, the expansion of $(\frac{n+1}{n}\cdot x)^k$ in the Gram basis with parameter $n+1$ is given by

$$C_k\cdot\phi_k + C_{k-1}\cdot\phi_{k-1} + \ldots + C_0\cdot\phi_0,$$

where $C_k \lesssim \left(\frac{n+1}{n}\right)^k \cdot \frac{1}{(2(n+1))^k} \cdot \sqrt{\frac{(n+k+1)!}{(n-k+1)!}} \lesssim \frac{1}{(2(n+1))^k} \cdot \sqrt{\frac{(n+k+1)!}{(n-k+1)!}}$. By Equation 4 and Cauchy-Schwarz, $\max_{D^{\text{in}}_{n+1}} \phi_k \lesssim 1$ and the corollary follows by taking the approximation $\sum_{i=0}^{k-1} C_i \cdot \phi_i$. ◀

Comparison to the uniform approach

Newman and Rivlin [13], and Sachdeva and Vishnoi [16] showed that any degree $\leq k-1$ uniform approximation to the monomial x^k over $[-1,1]$ will have error 2^{-k+1} (and that this is tight). For large enough values of n and k, the upper bound in the statement of Corollary 8 is smaller. In Section 7 we see that the bound 2^{-k+1} is insufficient to get a non-trivial indistinguishability upper bound for k close to n.

5 Hardness of discrete monomial approximations

▶ **Corollary 9.** *Fix integers $k < n$. Any degree at most $k-1$ polynomial approximation to the monomial x^k must have error (pointwise over D_n^{out}) at least:*

$$(2n)^{-k} \cdot \sqrt{\frac{(n+k)!}{(n-k)!}}$$

The proof of the main corollary of this section appears at the end of this section after two lemmas have been established.

▶ **Lemma 10.** *Let p be a degree k polynomial with degree k coefficient C in the discrete Chebyshev basis with parameter n. Then, any degree $k-1$ approximating polynomial will have error at least $|C|$ at some point over D_n^{in}.*

Proof. Let q be any degree at most $k-1$ polynomial. Let c_i be the degree i coefficient in the discrete Chebyshev representation of $p-q$, and note that because the degree of q is at most $k-1$, we have that $c_k = C$. By orthogonality of the discrete Chebyshev polynomials, and Equations 3 and 4,

$$\mathbb{E}_{t\sim D^{\text{in}}_n}[(p(t)-q(t))^2] = c_0^2 + \sum_{d=1}^{k} (c_d)^2\, \mathbb{E}_{t\sim D^{\text{in}}_n}[\phi_d(t)^2] \geq c_k^2 = C^2.$$

It follows that the approximation error $|p(t)-q(t)|$ must exceed $|C|$ for some $t \in D^{\text{in}}_n$. ◀

▶ **Lemma 11.** *Let p be a degree k polynomial such that for any degree at most $k-1$ polynomial q, $\max_{t\in D_n^{in}}\{|p(t)-q(t)|\} \geq \epsilon$. Then, for any degree $k-1$ polynomial q', $\max_{t\in D_n^{out}}\{|p(t)-q'(t)|\} \geq \epsilon$.*

Proof. We consider the contrapositive and show that existence of a degree at most $k-1$ polynomial $\tilde{p}$ for p over D^{out}_n with error at most ϵ implies an approximation for p over D^{in}_n with the same degree and error parameters. We have that $\tilde{p}(x+1/n)$ is a degree $k-1$ $\epsilon-$approximation of $p(1+1/n)$ over $\{x - \frac{1}{n} : x \in D^{\text{out}}_n\} \supset D^{\text{in}}_n$, where the set relation follows from Equation 2. Our approximation is then $\tilde{p}(t+1/n) + p(t) - p(t+1/n)$, which is degree $k-1$ because $p(t) - p(t+1/n)$ is degree $k-1$. We have that $\max_{t\in D^{\text{in}}_n} |\,(\tilde{p}(t+1/n) + p(t) - p(t+1/n)) - p(t)| = \max_{t\in D^{\text{in}}_n} |\tilde{p}(t+1/n) - p(t+1/n)| \leq \epsilon$. ◀

Proof of Corollary 9. Lemma 7 and Lemma 10 imply that any degree at most $k-1$ polynomial approximation to x^k must have error at least $(2n)^{-k} \cdot \sqrt{\frac{(n+k)!}{(n-k)!}}$ over D^{in}_n. The corollary then follows from Lemma 11. ◀

6 Symmetrization and duality

Our main upper and lower bounds will be justified by reducing to an approximation theoretic question using a linear programming duality relation.

▷ Claim 12 (see, for example, Theorem 1.2 in [2]). $\widetilde{\deg}_{\epsilon/2}(F) \geq k$ if and only if there exists a pair of perfectly k-wise indistinguishable distributions μ, ν over $\{0,1\}^n$ such that $\mathbb{E}_{X\sim\mu}[F(X)] - \mathbb{E}_{Y\sim\nu}[F(Y)] \geq \epsilon$.

We are interested in Boolean functions as statistical tests that witness k bits of a sample from a distribution. To this end, let Q_w denote the function on $\{0,1\}^k$ that outputs 1 if and only if the Hamming weight is exactly w. We also use $Q_w(x|_S)$ to represent Q_w when evaluated on a string of n bits, where $x|_S$ is the restriction of the n bits to the k indices in the set $S \subseteq [n]$.

▶ **Fact 13.** *Let $S \subseteq [n]$ be any set of size k. There exists a univariate polynomial p_w of degree at most k such that the following holds. For all $t \in D_n^{out}$, $p_w(t) = \mathbb{E}_Z[Q_w(Z|_S)]$ where Z is a uniformly random string of n bits conditioned on having Hamming weight $\phi^{-1}(t) = (1-t)n/2 \in \{0,1,\ldots,n\}$.*

Proof. This statement is a simple extension of Minsky and Papert's classic symmetrization technique [11] and also appears in [3]; we reproduce the proof here for convenience. Minsky and Papert showed that for any polynomial $P\colon \{0,1\}^n \to \mathbb{R}$, there exists a univariate polynomial p of degree at most the total degree of P, such that for all $i \in \{0,\ldots,n\}$, $p(i) = \mathbb{E}_{|x|=i}[P(x)]$. Apply this result to $P(x) = Q_w(x|_S)$ and let $p_w(t) = p(\phi^{-1}(t)) = p\left((1-t)n/2\right)$. The fact then follows from the observation that the total degree of $Q_w(x|_S)$ is at most k, since this function is a k-junta. ◀

▶ **Corollary 14.** *Suppose that for all degree $\leq k-1$ polynomials q we have that $\max_{t\in D_n^{out}}\{|p_w(t) - q|\} \geq \epsilon$. Then, $\widetilde{\deg}_\epsilon(Q_w(x|_S)) \geq k-1$.*

Proof. We prove the contrapositive. Let $\tilde{Q}$ be a degree at most $k-1$ approximation to $Q_w(x|_S)$ over $\{0,1\}^n$ with pointwise error strictly less than ϵ. Taking the Minsky-Papert symmetrization of $\tilde{Q}$ in conjunction with a scale and shift, yields a univariate polynomial q of degree at most $k-1$ such that $\max_{t\in D_n^{\text{out}}}\{|p_w(t) - q|\} < \epsilon$. ◀

6.1 Properties of p_w

The value $p_w(t)$ is a probability for every $t \in D_n^{\text{out}}$. Moreover, this probability must equal zero when the Hamming weight of Z is less than w or greater than $n-k+w$. Therefore p_w has k distinct zeros at the points $Z_w = Z_- \cup Z_+$, where

$$Z_- = \{-1 + 2h/n : h = 0, ..., k-w-1\}, \quad Z_+ = \{1 - 2h/n : h = 0, ..., w-1\}. \tag{13}$$

and so p_w must have the form

$$p_w(t) = C_w \cdot \prod_{z\in Z_w} (t - z) \tag{14}$$

for some C_w that does not depend on t.

▷ Claim 15. The coefficient C_w on the highest degree term of p_w in the monomial basis has absolute value:

$$\frac{\binom{k}{w}\binom{n-k}{\frac{1}{2}(n-k)}}{\binom{n}{\frac{1}{2}(n-k+2w)}} \cdot \frac{n^k \cdot (n-k)!!^2}{(n-k+2w)!! \cdot (n-2w+k)!!}$$

Proof. The polynomial p_w is of degree k with all of its zeroes lying in Z_w. We evaluate p_w at a point t' which is necessarily outside of Z_w and thus not a zero of p_w. We set:

$$t' := \frac{1}{2}\left(\max\{Z_-\} + \min\{Z_+\}\right) = \frac{k-2w}{n}$$

To evaluate $p_w(t')$, we use that the value $p_w(t')$ is the probability that $Q_w(x|_S)$ accepts given that x is chosen uniformly at random, conditioned on the event that the Hamming weight of x is exactly $\phi^{-1}(t') = \frac{1}{2}(n-k+2w)$.

$$\Pr\left[Q_w(x|_S) = 1 : |x| = \frac{1}{2}(n-k+2w)\right] = p_w(t') = C_w \cdot \prod_{z\in Z_w}(t'-z),$$

from which it follows that

$$C_w = \frac{\binom{k}{w}\binom{n-k}{\frac{1}{2}(n-k)}}{\binom{n}{\frac{1}{2}(n-k+2w)} \cdot \prod_{z\in Z_w}(t'-z)}$$

We have that:

$$\begin{aligned}\prod_{z\in Z_w}(t'-z) &= \prod_{z\in Z_-}(t'-z)\prod_{z\in Z_+}(t'-z)\\ &= (-1)^{|Z_+|}\prod_{z\in Z_+}(z-t')^2 \prod_{z\in Z_-:-z\notin Z_+}(t'-z),\end{aligned}$$

where the final equality assumes that $w \le k/2$. This is without loss of generality; when $w > k/2$, the same calculation holds with the roles of Z_+ and Z_- reversed. From this we compute that:

$$\begin{aligned}\frac{1}{|\prod_{z\in Z_w}(t'-z)|} &= \frac{1}{(1-\frac{k-2}{n})^2\cdot(1-\frac{k-2}{n}+\frac{2}{n})^2\cdot\ldots\cdot(1-\frac{k-2}{n}+\frac{2(w-1)}{n})^2\cdot\prod_{z\in Z_-:-z\notin Z_+}(t'-z)}\\ &= \frac{n^{2w}\cdot(n-k)!!^2}{(n-k+2w)!!^2\cdot\prod_{z\in Z_-:-z\notin Z_+}(t'-z)}\\ &= \frac{n^{2w}\cdot(n-k)!!^2}{(n-k+2w)!!^2\cdot\prod_{i=0}^{k-2w-1}(\frac{k-2w}{n}+1-\frac{2i}{n})}\\ &= \frac{n^{k}\cdot(n-k)!!^2}{(n-k+2w)!!\cdot(n+k-2w)!!},\end{aligned}$$

from which the claim follows. ◁

We also will need the following fact, which is justified in the Appendix.

▶ **Lemma 16.** *The value C_w is maximized when $w = k/2$; in particular with*

$$C_{k/2} = \frac{\binom{k}{k/2}\binom{n-k}{(n-k)/2}}{\binom{n}{n/2}} \cdot \frac{n^k\cdot(n-k)!!^2}{n!!^2}$$

7 Upper bound

We begin by providing an upper bound on the distinguishing advantage of a given Q_w test.

▶ **Lemma 17.** *For any $w = 0,\ldots,k$ and pair of $(k-1)$-wise indistinguishable distributions, the function Q_w reconstructs with advantage ϵ, satisfying:*

$$\epsilon \lesssim \frac{(n-k)^{\frac{n-k}{2}}\cdot(n+k)^{\frac{n+k}{2}}}{2^k\cdot n^n}$$

Proof. By Corollary 8, there exists a degree $k-1$ polynomial approximation to p_w over D_n^{out} with error

$$\lesssim C_w \cdot \frac{1}{(2(n+1))^k} \cdot \sqrt{\frac{(n+k+1)!}{(n-k+1)!}} \tag{15}$$

By Claim 15 and Lemma 16 this is upper bounded by (up to poly(n) factors):

$$\frac{\binom{k}{k/2}\binom{n-k}{(n-k)/2}}{\binom{n}{n/2}} \cdot \frac{n^k \cdot (n-k)!!^2}{n!!^2} \cdot \frac{1}{(2(n+1))^k} \cdot \sqrt{\frac{(n+k+1)!}{(n-k+1)!}}$$

$$\begin{aligned}
&\lesssim \frac{k! \cdot (n-k)! \cdot n^k}{(k/2)!^2 \cdot n! \cdot 2^k} \cdot \frac{1}{(2n+2)^k} \cdot \sqrt{\frac{(n+k+1)!}{(n-k+1)!}} \\
&\lesssim \frac{(n-k)^{n-k} \cdot (2e)^k}{n^n} \cdot \frac{n^k}{2^k} \cdot \frac{1}{(2n+2)^k} \cdot \frac{(n+k+1)^{(n+k+1)/2)}}{e^k \cdot (n-k+1)^{(n-k+1)/2}} \\
&\lesssim \frac{(n-k)^{(n-k-1)/2} \cdot (n+k+1)^{(n+k+1)/2}}{2^k \cdot n^n} \\
&\lesssim \frac{(n-k)^{(n-k)/2} \cdot (n+k)^{(n+k)/2}}{2^k \cdot n^n},
\end{aligned}$$

where we have used Equations 11, 10, and 9 to bound the central binomial coefficient, the factorial, and the double factorial, respectively. ◀

Comparison to the uniform approach

We saw in Section 4 that any *uniform* approximation to the monomial would have error 2^{-k+1}. Substituting that bound into Equation 15 and carrying out the same calculation would yield an upper bound of $\frac{e^k \cdot (n-k)^{n-k}}{2^k \cdot n^{n-k}}$, which for $k \approx n$ is $\gtrsim (e/2)^n$. Because any distinguishing advantage must be at most 1, this is a vacuous bound.

7.1 Proof of upper bound: Theorem 1

▶ **Theorem 18.** *For any pair of $(k-1)$-wise indistinguishable distributions μ, ν over $\{0,1\}^n$, the statistical distance ϵ between $\mu|_k$ and $\nu|_k$ satisfies:*

$$\epsilon \lesssim \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}$$

Proof. Let T be a general distinguisher on k inputs. By Facts 5 and 6, T can be assumed to be a symmetric Boolean-valued function and has the representation $T = \sum_{w=0}^{k} b_w \cdot Q_w$ where each of the b_w is either 0 or 1. We bound the distinguishing advantage as follows. Recalling that μ and ν are $k-1$-indistinguishable symmetric distributions over $\{0,1\}^n$, for any set $S \subseteq [n]$ of size k we have:

$$\begin{aligned}
\mathbb{E}_{X\sim\mu}[T(X|_S)] - \mathbb{E}_{Y\sim\nu}[T(Y|_S)] &= \sum_{w=0}^{k} b_w \big(\mathbb{E}[Q_w(X|_S)] - \mathbb{E}[Q_w(Y|_S)]\big) \\
&\leq \sum_{w=0}^{k} \big|\mathbb{E}[Q_w(X|_S)] - \mathbb{E}[Q_w(Y|_S)]\big| \\
&\leq (k+1) \cdot \max_{w=0,\ldots,k} \big|\mathbb{E}[p_w(\phi(|X|)] - \mathbb{E}[p_w(\phi(|Y|))]\big| \\
&\lesssim \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n},
\end{aligned}$$

where the final upper bound is from Lemma 17. ◀

8 Proof of lower bound: Theorem 2

▶ **Theorem 19.** *Let $Q_{k/2}$ be the statistical test over k bits that accepts if and only if the observed Hamming weight is $k/2$. There exists a pair of $(k-1)$-wise indistinguishable distributions X, Y such that the reconstruction advantage of $Q_{k/2}$ is at least ϵ, satisfying:*

$$\epsilon \gtrsim \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}.$$

Proof. By Claim 12 and Corollary 14, it suffices to show that any degree $k-1$ polynomial approximation to $p_{k/2}$ over D_n^{out} must have error at least ϵ. Lemma 11 reduces the problem further to proving hardness of approximation of $p_{k/2}$ over D_n^{in}. From Claim 15 and Lemma 7, the coefficient on the degree k term of the discrete Chebyshev representation of $p_{k/2}$ is

$$\gtrsim \frac{\binom{k}{k/2}\binom{n-k}{(n-k)/2}}{\binom{n}{n/2}} \cdot \frac{n^k \cdot (n-k)!!^2}{n!!^2} \cdot \frac{1}{(2n)^k} \cdot \sqrt{\frac{(n+k)!}{(n-k)!}},$$

which is $\gtrsim \epsilon$, by trivially applying the bounds for the central binomial coefficient, factorial, and double factorial in Section 2. The theorem then follows by applying Lemma 10. ◀

9 Future research

It would be worthwhile to explicitly construct the distributions of Theorem 2 (as done in [10] to match the non-constructive bounds in [9]). We also ask whether similar methods can be extended to work in the full setting of [3, 8] in which there may be a gap between the indistinguishability and reconstruction parameters or in the context of distributions μ, ν over Σ^n, where Σ is an alphabet of size larger than 2. Finally, we observe the interplay between results stemming from classical approximation theory (often featuring Chebyshev polynomials) and quantum algorithms (notably in the pair of papers [18, 7]) and ask whether our results can be recovered with a quantum argument.

References

1 RW Barnard, Germund Dahlquist, K Pearce, Lothar Reichel, and KC Richards. Gram polynomials and the kummer function. *Journal of approximation theory*, 94(1):128–143, 1998.

2 Andrej Bogdanov, Yuval Ishai, Emanuele Viola, and Christopher Williamson. Bounded indistinguishability and the complexity of recovering secrets. In *CRYPTO*, 2016.

3 Andrej Bogdanov, Nikhil S Mande, Justin Thaler, and Christopher Williamson. Approximate degree, secret sharing, and concentration phenomena. *RANDOM*, 2019.

4 Andrej Bogdanov and Christopher Williamson. Approximate bounded indistinguishability. In *ICALP*, 2017.

5 H. Buhrman, R. Cleve, R. de Wolf, and C. Zalka. Bounds for small-error and zero-error quantum algorithms. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, 1999.

6 Mark Bun and Justin Thaler. Dual lower bounds for approximate degree and markov–bernstein inequalities. *Information and Computation*, 243:2–25, 2015.

7 Ronald de Wolf. A note on quantum algorithms and the minimal degree of epsilon-error polynomials for symmetric functions. *arXiv preprint arXiv:0802.1816*, 2008.

8 Xuangui Huang and Emanuele Viola. Approximate degree-weight and indistinguishability. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 26, page 85, 2019.

9 Matthias Krause and Hans Ulrich Simon. Determining the optimal contrast for secret sharing schemes in visual cryptography. In Gaston H. Gonnet and Alfredo Viola, editors, *LATIN 2000: Theoretical Informatics*, volume 1776 of *Lecture Notes in Computer Science*, pages 280–291. Springer Berlin Heidelberg, 2000.

10 Christian Kuhlmann and Hans Ulrich Simon. Construction of visual secret sharing schemes with almost optimal contrast. In *Symposium on Discrete Algorithms: Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms*, volume 9 (11), pages 263–272, 2000.

11 Marvin Minsky and Seymour Papert. *Perceptrons*. MIT Press, Cambridge, MA, 1969.

12 Moni Naor and Adi Shamir. Visual cryptography. In *Advances in Cryptology – EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 1994.

13 DJ Newman and TJ Rivlin. Approximation of monomials by lower degree polynomials. *aequationes mathematicae*, 14(3):451–455, 1976.

14 Noam Nisan and Mario Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994.

15 Ramamohan Paturi. On the degree of polynomials that approximate symmetric boolean functions (preliminary version). In *ACM Symp. on the Theory of Computing (STOC)*, pages 468–474, 1992.

16 Sushant Sachdeva and Nisheeth Vishnoi. Approximation theory and the design of fast algorithms. *arXiv preprint*, 2013. `arXiv:1309.4882`.

17 Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *40th ACM Symp. on the Theory of Computing (STOC)*, pages 85–94, 2008.

18 Alexander A Sherstov. Approximate inclusion-exclusion for arbitrary symmetric functions. *Computational Complexity*, 18(2):219–247, 2009.

A A technical claim

▷ Claim 20. Let $v = \prod_{i=1}^{k} \frac{(i-1/2)(i+1/2)}{i^2}$. Then,

$$\frac{3}{4k^2} \leq v \leq 1$$

Proof. We have that $v = \prod_{i=1}^{k} \left(1 - \frac{1}{4i^2}\right)$, which is a product of numbers less than 1 and justifies the upper bound. For the lower bound, we have:

$$\begin{aligned}
\frac{1}{v} &= \prod_{i=1}^{k} \frac{4i^2}{4i^2-1} \\
&= \frac{4}{3} \cdot \frac{16}{15} \cdot \ldots \cdot \frac{4k^2}{4k^2-1} \\
&\leq \frac{4}{3} \cdot \frac{16}{4} \cdot \ldots \cdot \frac{4k^2}{4(k-1)^2} \\
&= \frac{4k^2}{3}.
\end{aligned}$$

◁

B Proof of Lemma 16

Proof. We find the maximising value of C_w by expanding the expression for C_w and removing terms that do not depend on w:

$$\begin{aligned}
\arg\max_w C_w &= \arg\max_w \frac{\binom{k}{w}\binom{n-k}{\frac{1}{2}(n-k)}}{\binom{n}{\frac{1}{2}(n-k+2w)}} \cdot \frac{n^k \cdot (n-k)!!^2}{(n-k+2w)!! \cdot (n-2w+k)!!} \\
&= \arg\max_w \frac{\binom{k}{w}}{\binom{n}{\frac{1}{2}(n-k+2w)}} \cdot \frac{1}{(n-k+2w)!! \cdot (n-2w+k)!!} \\
&= \arg\max_w \frac{k! \cdot (\frac{n-k+2w}{2})! \cdot (n-\frac{n-k+2w}{2})!}{w! \cdot (k-w)! \cdot n! \cdot (n-k+2w)!! \cdot (n-2w+k)!!} \\
&= \arg\max_w \frac{(\frac{n-k+2w}{2})! \cdot (\frac{n+k-2w}{2})!}{w! \cdot (k-w)! \cdot (n-k+2w)!! \cdot (n-2w+k)!!} \\
&= \arg\max_w \frac{(\frac{n-k+2w}{2})! \cdot (\frac{n+k-2w}{2})!}{w! \cdot (k-w)! \cdot 2^n \cdot (\frac{n-k+2w}{2})! \cdot (\frac{n-2w+k}{2})!} \\
&= \arg\max_w \frac{1}{w! \cdot (k-w)!} = k/2
\end{aligned}$$

◀

C Proof of Corollary 3

In this section we justify Corollary 3. The proof will rely on two technical lemmas which are presented prior to the final proof of the corollary.

▶ **Lemma 21.** *Let $g = g(\epsilon)$ be defined over $[0, 0.4)$, where g is:*

$$\frac{(1.6+\epsilon)^{0.8+\epsilon/2} \cdot (0.4-\epsilon)^{0.2-\epsilon/2}}{2^{0.6+\epsilon}}$$

Then, for any positive ϵ in the domain of g, we have $g(\epsilon) > 4/5$.

Proof. It suffices to show: (a) that g is minimised over its domain at 0 and $g(0) = 4/5$ and (b) that $g' > 0$ for all positive ϵ. Computing the derivative, we have:

$$g'(\epsilon) = (1/5) \cdot \left(2^{-1.6-\epsilon} \cdot (2-5\epsilon)^{0.2-\epsilon/2} \cdot (8+5\epsilon)^{0.8+\epsilon/2} \cdot (\ln(1.6+\epsilon) - \ln(1.6-4\epsilon))\right)$$

The statement (a) is justified by checking that $4/5 = g(0) \leq \lim_{\epsilon \to 0.4} g(\epsilon)$ and checking that the derivative is zero only at $\epsilon = 0$. Statement (b) follows from observing that g' is a product of exponential terms that are always positive and the factor $\ln(1.6+\epsilon) - \ln(1.6-4\epsilon)$, which is also always positive for positive ϵ because the natural log is an increasing function. ◀

▶ **Lemma 22.** *Let f_n denote the function of k in the lower bound of Theorem 2, when n is sufficiently large and fixed. Similarly, let F_n be the function of k in the upper bound Theorem 1 when n is sufficiently large and fixed. Then, for any fixed $\epsilon > 0$, $F_n(0.6n) < f_n((0.6+\epsilon)n)$.*

Proof. Fix $\epsilon > 0$ and let δ be $g(\epsilon) - 4/5$. By Lemma 21, $\delta > 0$. The present lemma then follows immediately from the inequalities:

$$F_n(0.6n) = O(n^c) \cdot (4/5)^n \leq (4/5+\delta)^n \leq g^n(\epsilon) = f_n((0.6+\epsilon)n),$$

where the first equality is from Theorem 1, the first inequality is true for large enough n, the second inequality follows from Lemma 21, and the final equality is from the definition of f_n and Theorem 2. ◀

We now use Lemmas 21 and 22 to prove Corollary 3.

Proof. Fix arbitrary $\epsilon > 0$. In the game of Corollary 3, suppose that Player 1 chooses $\tilde{k}$, where

$$\tilde{k} > (0.6 + \epsilon) \cdot n.$$

Then, by Theorem 2 and Lemma 22, Player 2 will be able to choose a symmetric pair of perfectly $(\tilde{k} - 1)$-wise indistinguishable distributions where the $\tilde{k}$-wise reconstruction advantage is at least $f_n(\tilde{k}) > F_n(0.6n)$. Thus, the payoff of Player 1 will be strictly less than $1 - F_n(0.6n)$. This $\tilde{k}$ could not have been the optimal strategy for Player 1, since Player 1 can choose $0.6n$, where by Theorem 1, Player 2 achieves payoff at most $F_n(0.6n)$ and Player 1 gets payoff at least $1 - F_n(0.6n)$. Thus, the optimal strategy k^* cannot be chosen to be as large as $\tilde{k}$ and we have that

$$k^* - 0.6n \leq \epsilon n$$

We omit the proof of the complementary lower bound on k^* because it follows exactly the same structure as the upper bound. Together, these bounds imply that $|k^* - 0.6n| \leq \epsilon n$ for any arbitrarily small positive constant ϵ. ◀

D Proof of Corollary 4

Proof. Suppose that there exists an f that obeys the premises of the corollary, namely symmetry, no low-degree terms, and symmetry of Fourier coefficients. Define distributions μ, ν from f as follows (as in the method of [2]): $\mu(z) = 2 \cdot \max\{0, f(z)\}, \nu(z) = 2 \cdot \max\{0, -f(z)\}$. The total weight of each distribution is 1 because $\sum_z |f(z)| = 1$ and because by assumption $\hat{f}(\varnothing) = 0$, which implies that $\sum_z f(z) = 0$. Thus μ and ν are valid distributions. Next, observe that for every function $\chi_S := \prod_{i \in S} x_i$, the function χ_S has zero distinguishing advantage between μ and ν when $|S| \leq k - 1$:

$$\hat{f}(S) = 0 \implies \sum_z \chi_S(z) f(z) = 0 \implies \frac{1}{2}\left(\sum_z \chi_S(z)\mu(z) - \sum_z \chi_S(z)\nu(z)\right) = 0$$

$$\implies \mathbb{E}_{X \sim \mu}[\chi_S(X|_S)] - \mathbb{E}_{Y \sim \nu}[\chi_S(Y|_S)] = 0.$$

The set of all χ_S for $|S| \leq k - 1$ form a basis for all statistical tests on $k - 1$ bits; thus it follows immediately that μ and ν are perfectly $(k-1)$-wise indistinguishable. The symmetry of μ and ν follows straightforwardly from the symmetry of the Fourier coefficients of f.

We then have that for any $S \subseteq [n]$ where $|S| = k$,

$$\hat{f}(S) = \mathbb{E}[\chi_S \cdot f] = \frac{1}{2}\mathbb{E}[\chi_S \cdot (\mu - \nu)] = \mathbb{E}_{X \sim \mu}[\chi_S(X|_S)] - \mathbb{E}_{Y \sim \nu}[\chi_S(Y|_S)]$$

$$\leq O(n^c) \cdot \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}$$

The final inequality follows from Theorem 1 and fact that μ and ν obey the premises of that theorem, the implication that the statistical distance between any k-wise marginals of μ and ν is at most $O(n^c) \cdot \frac{(n-k)^{\frac{n-k}{2}} \cdot (n+k)^{\frac{n+k}{2}}}{2^k \cdot n^n}$, the fact that χ_S is a statistical test on k bits, and the fact that $\mathbb{E}_{X \sim \mu}[\chi_S(X|_S)] - \mathbb{E}_{Y \sim \nu}[\chi_S(Y|_S)]$ is precisely the reconstruction advantage of the statistical test χ_S. ◀